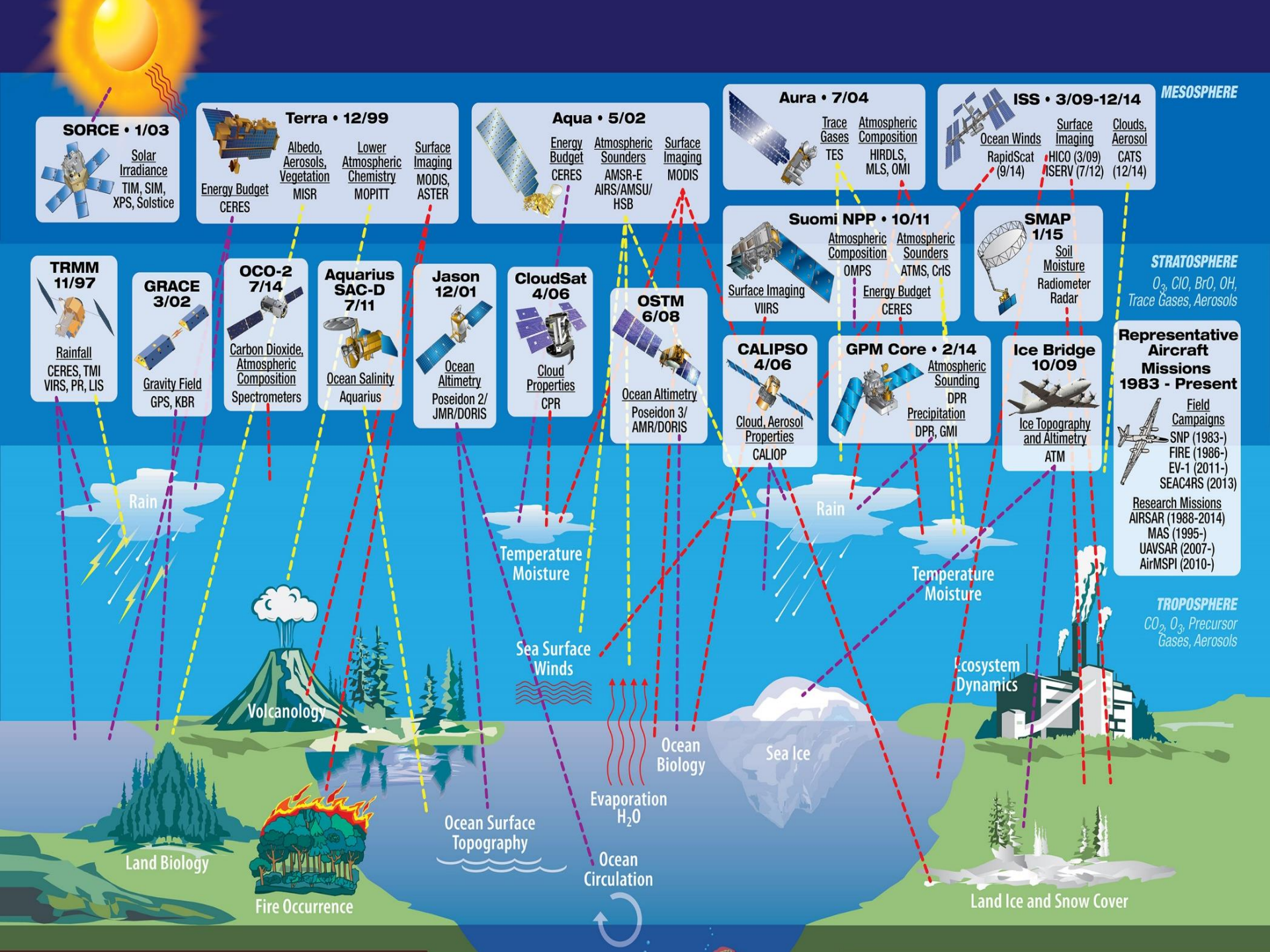
Lessons Learned while Exploring Cloud-Native Architectures for NASA EOSDIS Applications and Systems

Dan Pilone (dan@element84.com)
Brett McLaughlin (brett@element84.com)
Peter Plofchan (Peter.G.Plofchan@raytheon.com)

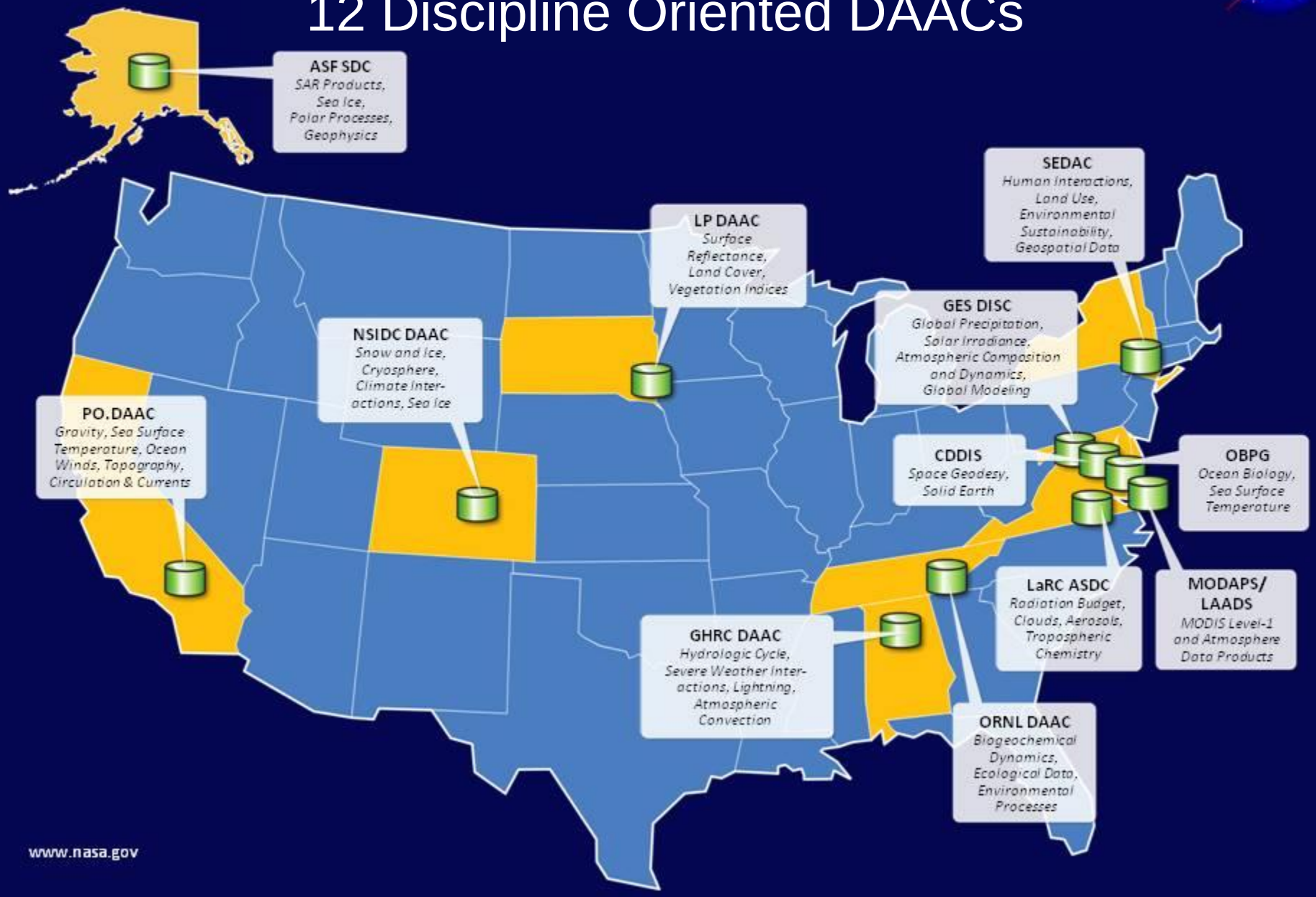
The material is based upon work supported by the National Aeronautics and Space Administration under Raytheon Contract Number NNG15HZ39C


Formulation
Implementation
Primary Ops
Extended Ops
Sentinel-6A/B
Earth Science Instruments on ISS:
RapidScat, CATS,
LIS, SAGE III (on ISS), TSIS-1, OCO-3, ECOSTRESS,
GEDI, CLARREO-PF
Landsat 9
PACE
NI-SAR
SWOT
TEMPO
JPSS-2 (NOAA)
RBI, OMPS-Limb
GRACE-FO (2)
ICESat-2
CYGNSS
ISS
SORCE,
TCTE (NOAA)
NISTAR, EPIC
(NOAA'S DSCOVR)
QuikSCAT
EO-1
Landsat 7
(USGS)
SMAP
Terra
Suomi NPP (NOAA)
Aqua
CloudSat
Landsat 8
(USGS)
CALIPSO
Aura
GRACE (2)
OSTM/Jason 2
(NOAA)
GPM
OCO-2


SORCE • 1/03
Solar Irradiance
TIM, SIM, XPS, Solstice
Terra • 12/99
Energy Budget
CERES
Albedo, Aerosols, Vegetation
MISR
Lower Atmospheric Chemistry
MOPITT
Surface Imaging
MODIS, ASTER
Aqua • 5/02
Energy Budget
CERES
Atmospheric Sounders
AMSR-E AIRS/AMSU/ HSB
Surface Imaging
MODIS
Aura • 7/04
Trace Gases
TES
Atmospheric Composition
HIRDLS, MLS, OMI
ISS • 3/09-12/14
Ocean Winds
RapidScat (9/14)
Surface Imaging
HICO (3/09) ISERV (7/12)
Clouds, Aerosol
CATS (12/14)
MESOSPHERE
TRMM 11/97
Rainfall
CERES, TMI VIRS, PR, LIS
GRACE 3/02
Gravity Field
GPS, KBR
OCO-2 7/14
Carbon Dioxide, Atmospheric Composition
Spectrometers
Aquarius SAC-D 7/11
Ocean Salinity
Aquarius
Jason 12/01
Ocean Altimetry
Poseidon 2/ JMR/DORIS
CloudSat 4/06
Cloud Properties
CPR
OSTM 6/08
Ocean Altimetry
Poseidon 3/ AMR/DORIS
Suomi NPP • 10/11
Surface Imaging
VIIRS
Atmospheric Composition
OMPS
Atmospheric Sounders
ATMS, CrIS
Energy Budget
CERES
SMAP 1/15
Soil Moisture
Radiometer Radar
STRATOSPHERE
O_3, ClO, BrO, OH, Trace Gases, Aerosols
CALIPSO 4/06
Cloud, Aerosol Properties
CALIOP
GPM Core • 2/14
Atmospheric Sounding
DPR
Precipitation
DPR, GMI
Ice Bridge 10/09
Ice Topography and Altimetry
ATM
Representative Aircraft Missions 1983 - Present
Field Campaigns
SNP (1983-)
FIRE (1986-)
EV-1 (2011-)
SEAC4RS (2013)
Research Missions
AIRSAR (1988-2014)
MAS (1995-)
UAVSAR (2007-)
AirMSPI (2010-)
TROPOSPHERE
CO_2, O_3, Precursor Gases, Aerosols
Rain
Temperature
Moisture
Rain
Temperature
Moisture
Sea Surface Winds
Volcanology
Ecosystem Dynamics
Ocean Biology
Sea Ice
Evaporation H_2O
Land Biology
Ocean Surface Topography
Ocean Circulation
Fire Occurrence
Land Ice and Snow Cover

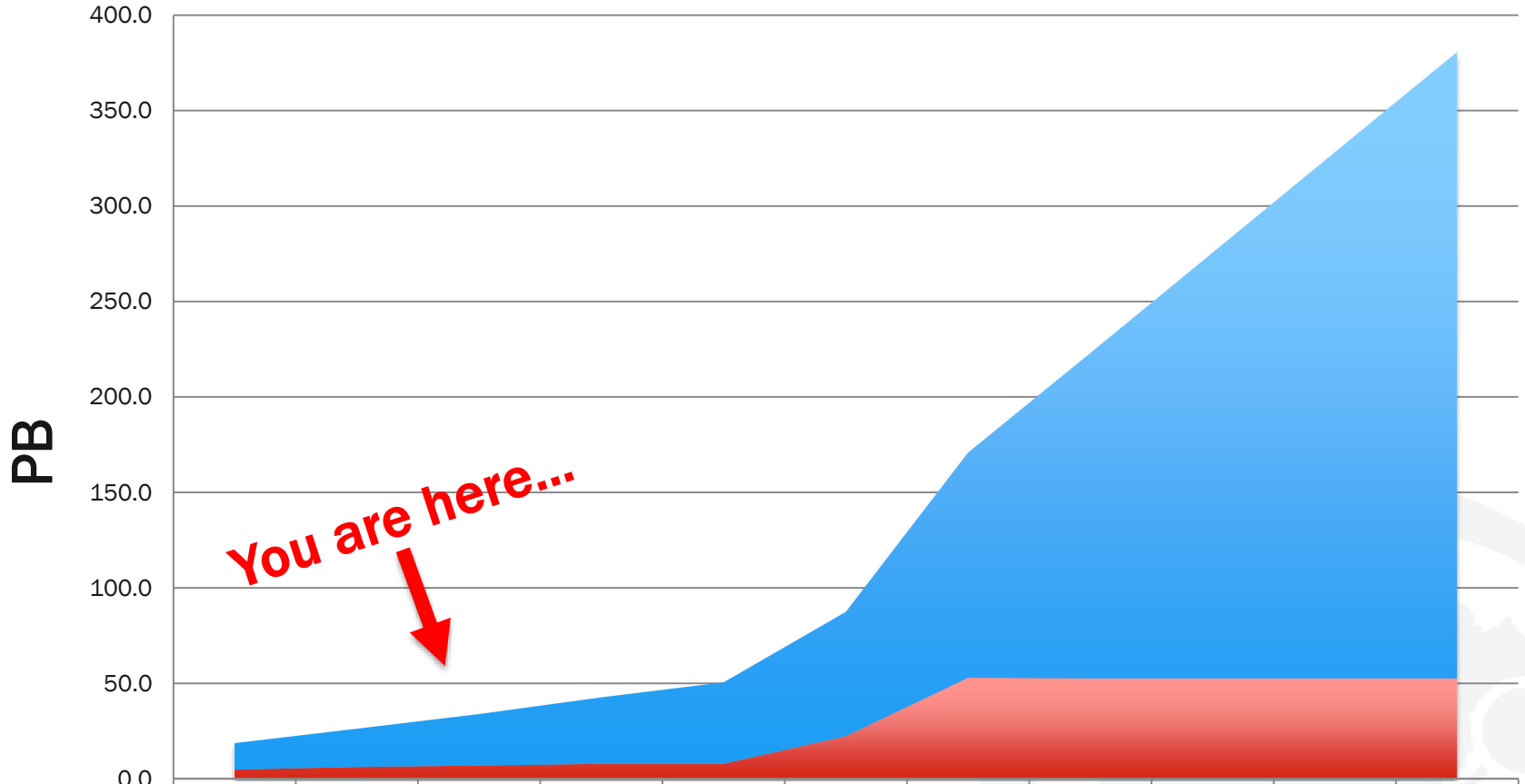
National Aeronautics and Space Administration

12 Discipline Oriented DAACs



www.nasa.gov

EOSDIS Archive Growth *Estimate* (Prime + Extended)



	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025
Cumulative Archive Size (PB)	13.8	20.0	27.0	34.8	42.7	65.0	118.0	170.5	223.1	275.6	328.2
Archive Growth Rate (PB)	4.9	6.2	7.0	7.9	7.9	22.4	52.9	52.6	52.6	52.6	52.6

EOSDIS

Lots of assumptions in this chart. Subject to change...

ExCEL Efforts and Project Prototypes

NGAP

NASA Compliant General Application Platform (NGAP), an operational, dev-ops, and sandbox AWS cloud based operating environment.

ASF WOS Prototype

AWS/NGAP Web Object Storage (WSO) prototyping large volumes of mission data dynamically between AWS S3, S3-IA, and Glacier object storage. Managed out of Alaska Satellite Facility

Earthdata Search Client to Cloud

NASA Earth Science data search by keyword and advanced filters such as time and space

Cumulus

Prototype addressing core EOSDIS capabilities including data ingest, archive, management, and distribution of large volumes of EOS data.

Getting Ready for NISAR (GRFN)

Integrated prototype of science product generation and delivery from a DAAC system focused on coupling ASF DAAC and JPL ARIA systems.

CATEES

Easy-to-use Python tools packaged to support EOSDIS cross-DAAC science workflows and analytics over large volumes of EOS data in AWS.

ECC to Cloud Study

Earth Code Collaborative (ECC) study to determine cloud ready capabilities to migrate into AWS/NGAP platform.

ExCEL Efforts and Project Prototypes Continued

GIBS in the Cloud

Migrating GIBS to the AWS/NGAP Cloud based on recommendations made in the "GIBS in the Cloud Study"

Earthdata Login to Cloud Study

Study to determine and recommend migrating the Earthdata Login into AWS/NGAP cloud environment

CMR to Cloud

Migration of the Common Metadata Repository, into the AWS/NGAP platform based on recommendations made in the CMR to Cloud study.

OPeNDAP/HDF Cloud Studies

Study to determine and recommend a cloud native integration of OPeNDAP accessing HDF5 and netCDF4 data on AWS/NGAP platform.

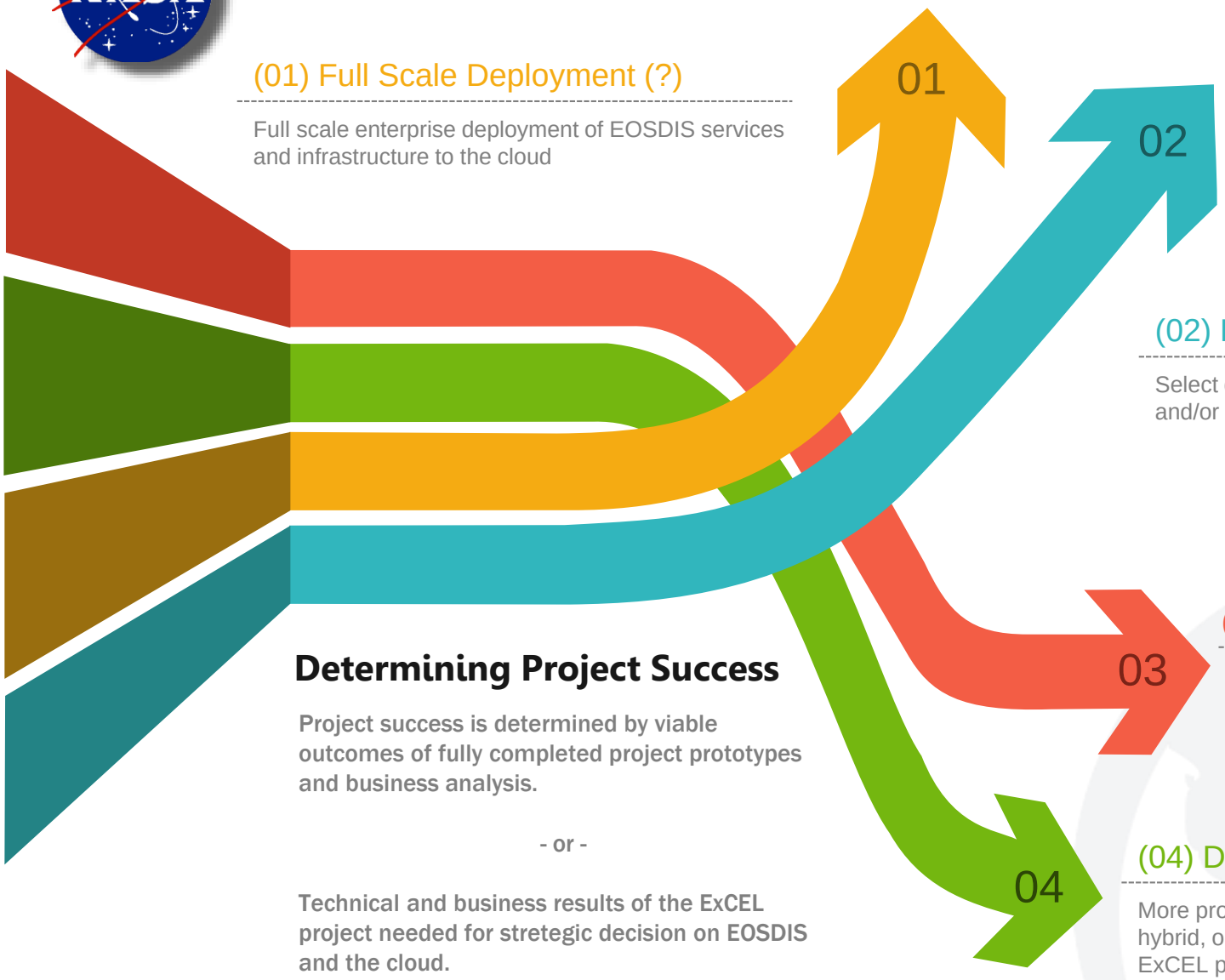
NEXUS

Prototype to accelerate end-user analysis of remote sensing data, highly parallel to better enable science discovery

Network Prototypes

Network prototypes to support to test security, monitoring, logging, and to perform R&D testing to support all ExCEL project prototypes.

ExCEL Go/No-Go



(01) Full Scale Deployment (?)

Full scale enterprise deployment of EOSDIS services and infrastructure to the cloud

(02) Partial Deployment (?)

Select deployment of EOSDIS services and/or infrastructure to the cloud

(03) Cloud Stand-down (?)

No EOSDIS services or infrastructure operationally migrated to the cloud

(04) Decision Point (?)

More prototyping required, or cloud hybrid, or other next steps based on ExCEL prototyping and business analysis results

Determining Project Success

Project success is determined by viable outcomes of fully completed project prototypes and business analysis.

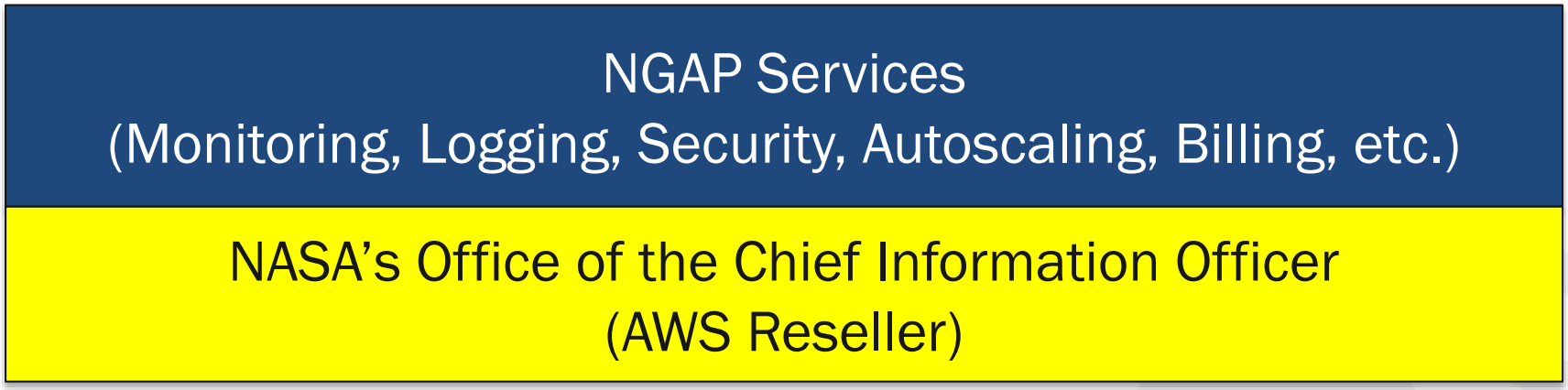
- or -

Technical and business results of the ExCEL project needed for stretegic decision on EOSDIS and the cloud.

What is NGAP?

NGAP is the **NASA General Application Platform**. It provides a cloud-based Platform-as-a-Service (PaaS) and Infrastructure-as-a-Service (IaaS) for ESDIS applications.

NGAP as a Platform



A Rough Look at Separation

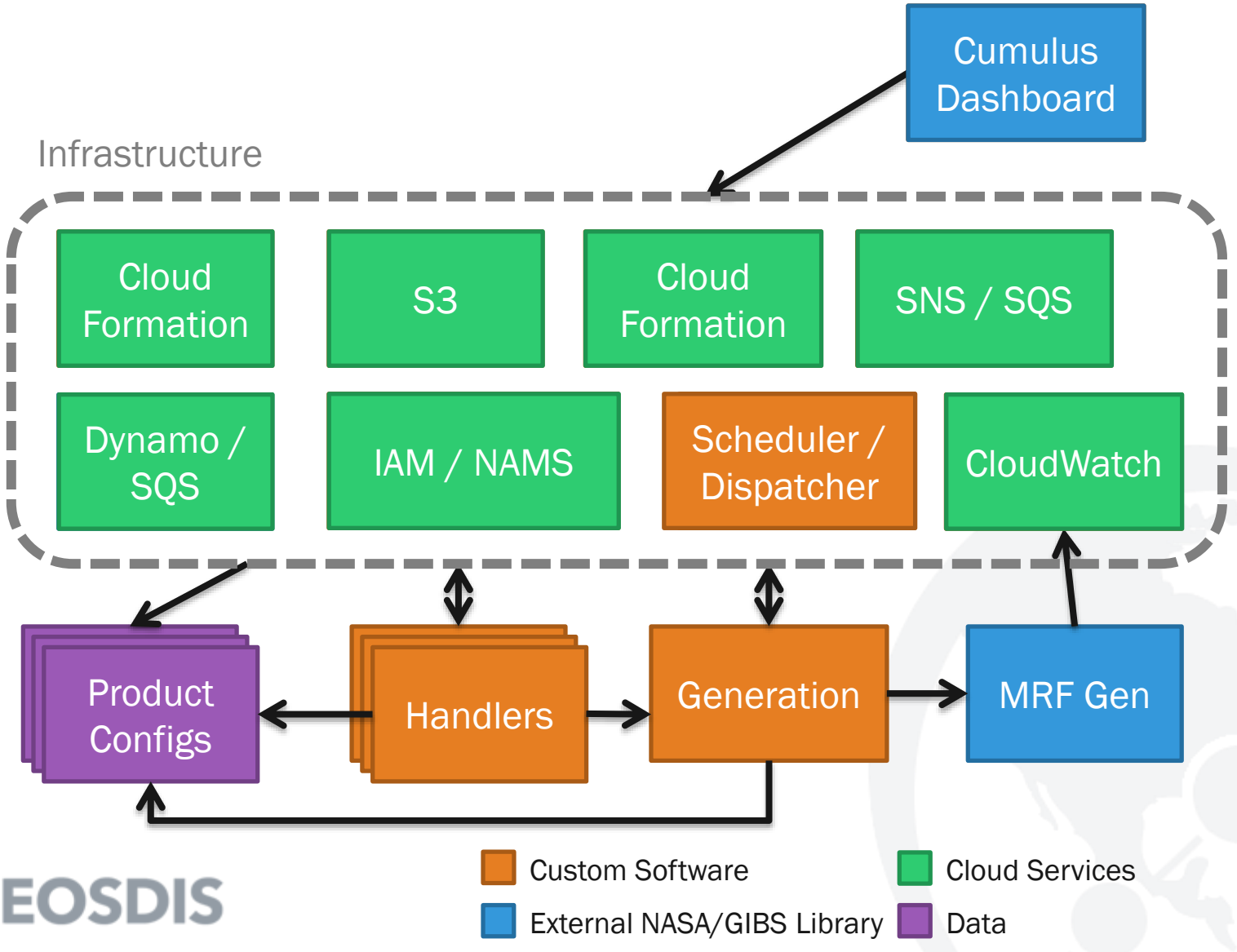
Lessons Learned

- Technical
- Psycho-Social
- Cost

Technical Lesson 1

ENABLE CLOUD NATIVE ARCHITECTURES BY STRONGLY PREFERRING CLOUD SERVICES

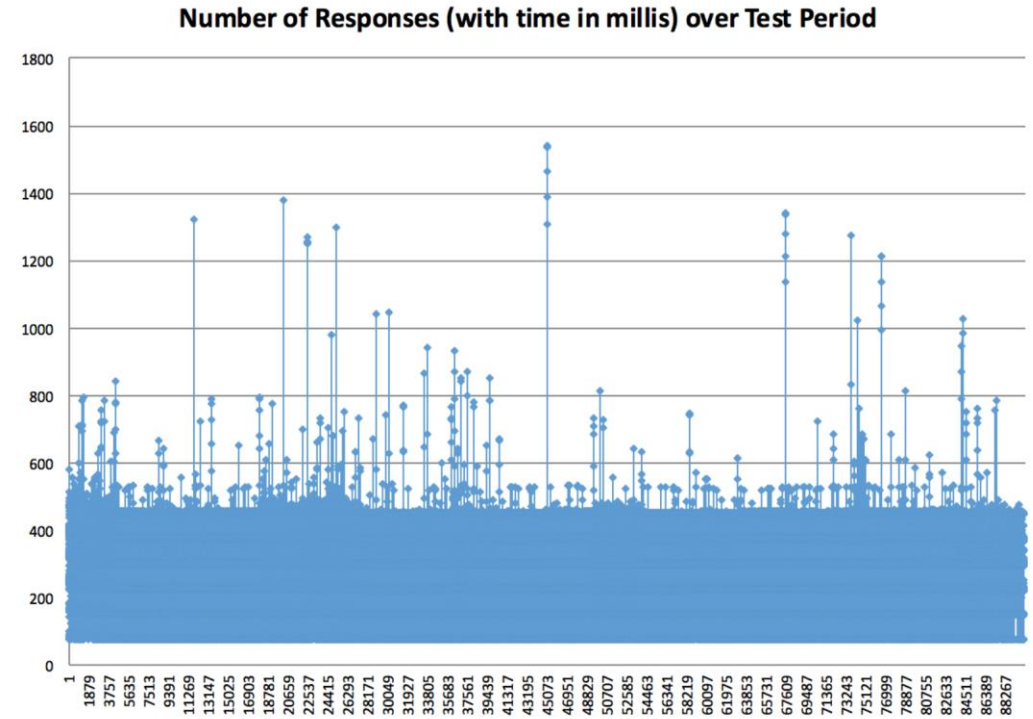
GIBS-in-the-Cloud Service Swap



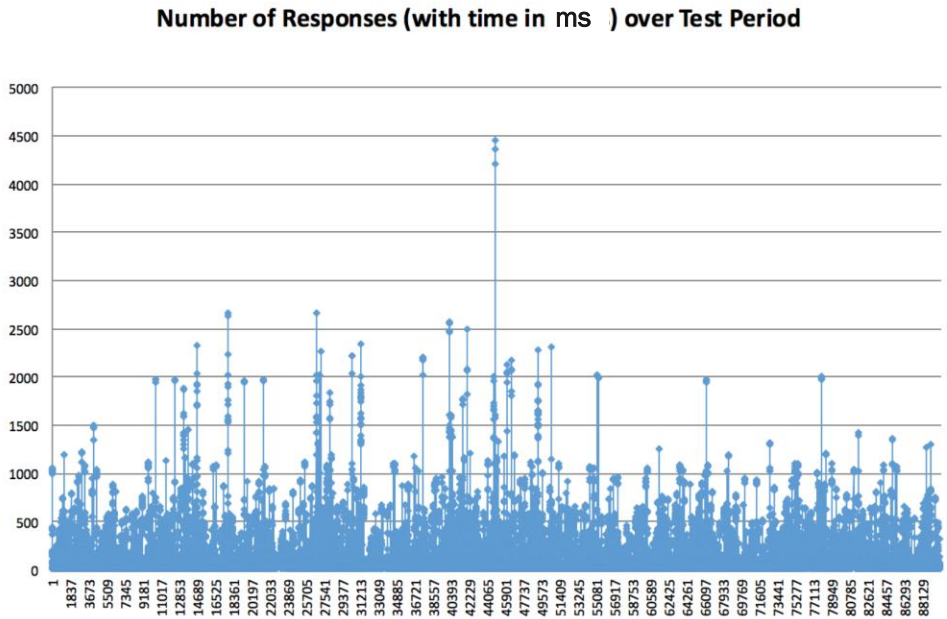
EOSDIS

Technical Lesson 2

AWS HAS VERY LOW INTERNAL LATENCY – BUT TRUST NOTHING.



On premises implementation showed consistent performance during load testing vs more sporadic latencies in AWS.

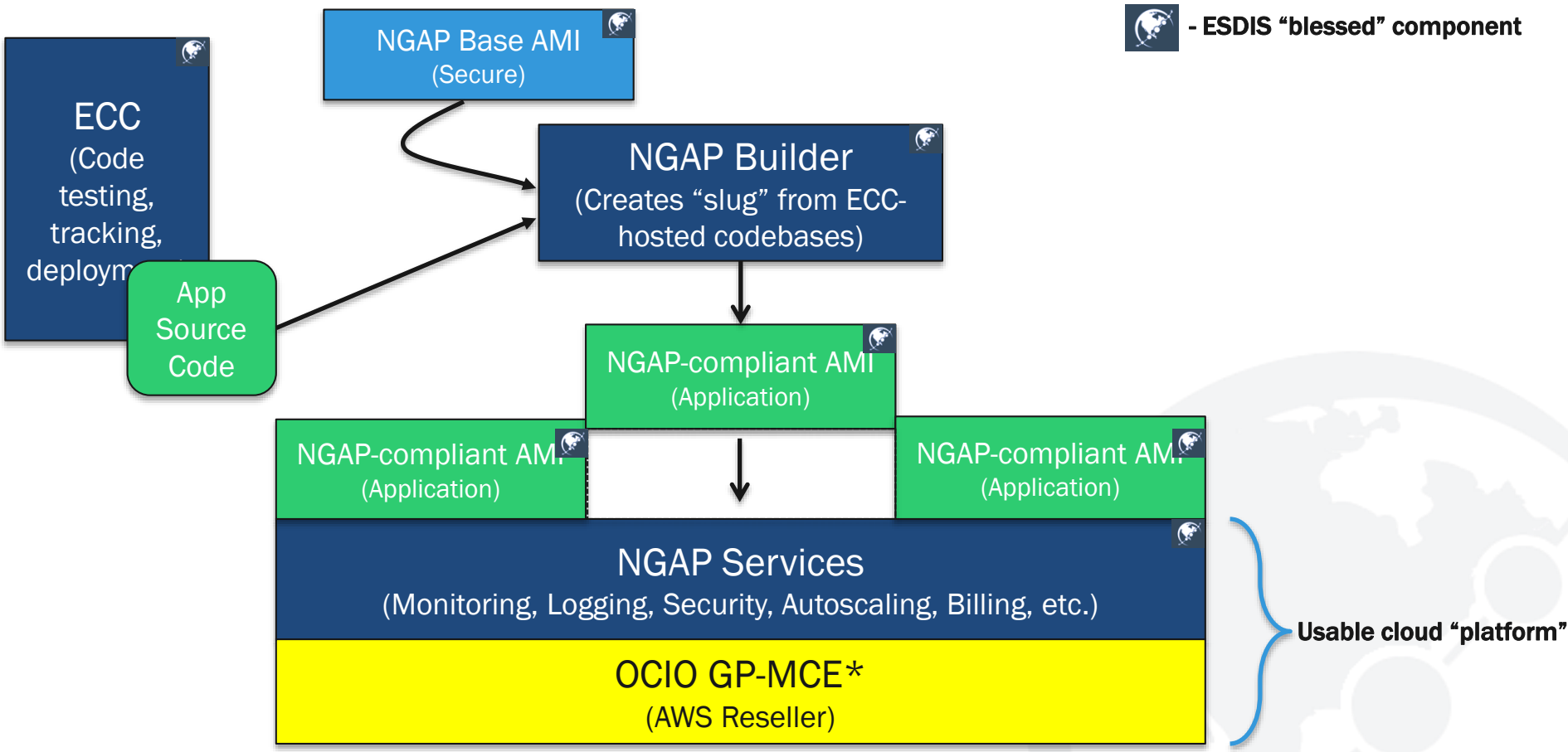


EOSDIS

Technical Lesson 3

INVOLVE SECURITY FROM THE VERY BEGINNING

Layer security thoughout the architecture



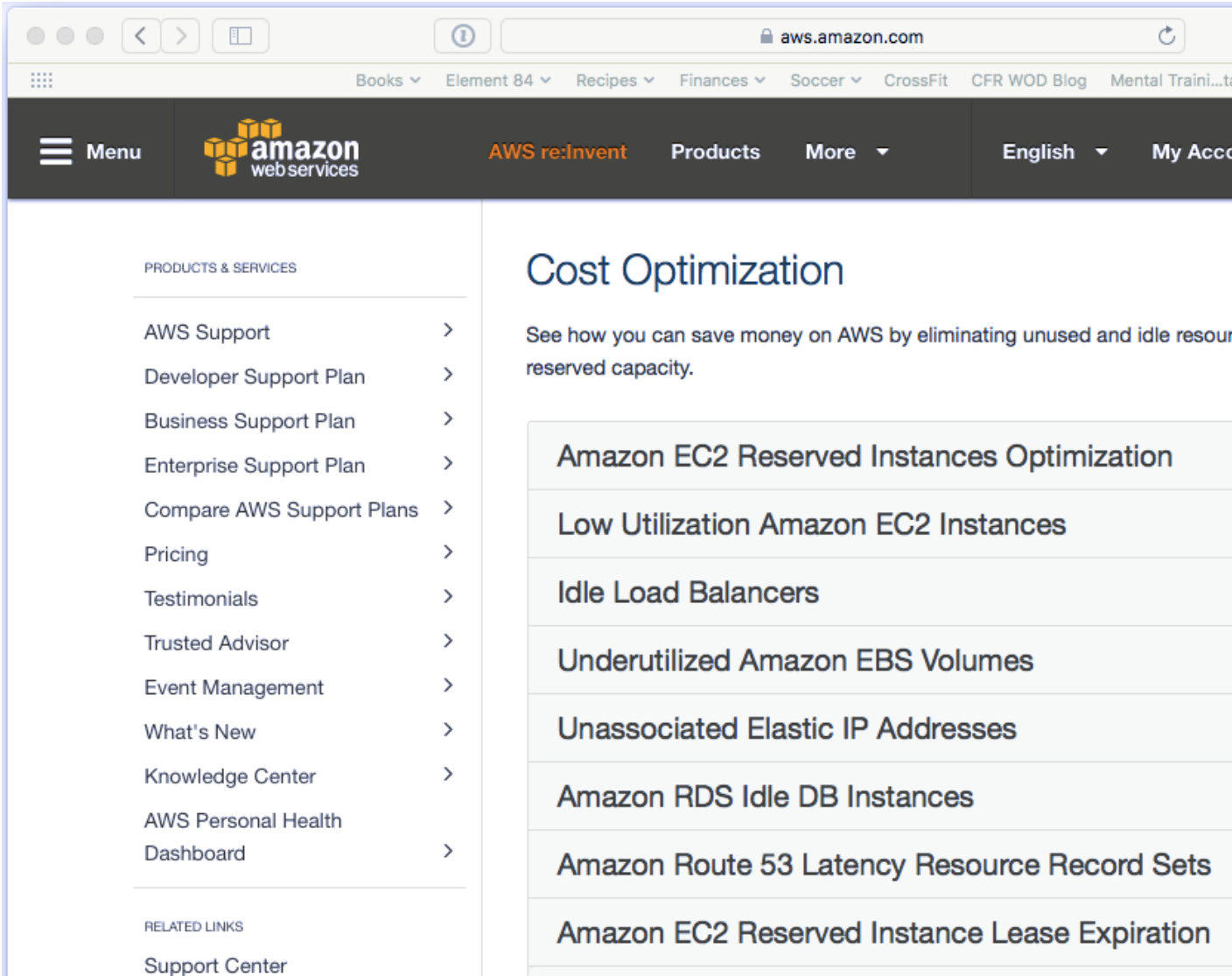
EOSDIS

*General Purpose Managed Compute Environment

Cost Lesson 1

MODELING TOTAL COST OF OWNERSHIP (TCO) IS EXTREMELY COMPLICATED

There's a lot to think about…



EOSDIS

The Big 4

1. EC2 Instances
   - More instances *running* = more cost
2. EBS Storage
   - More EBS* = more cost
3. Data Transfer
   - Notably: egress, egress, and egress
4. ELBs
   - More ELBs and more traffic = more cost

EOSDIS

** EBS storage has costs associated even when not in use by a running EC2 instance*

Also…

EOSDIS

May not use slides 23-26…

NGAP and Costs

COST CONCERNS IN NGAP

Planning over/around Auto-Scaling

- Autoscaling is available…
  - Most applications are setup in autoscaling groups
  - Autoscaling is 100% available within NGAP
- …but not completely automatic
  - NGAP disallows unbounded costs
  - NGAP favors planning over reaction

- **Takeaway**: *NGAP is more a hybrid than a true auto-scaling cloud solution*

NGAP is multi-region but not *all*-region

- NGAP exists in several regions…
  - us-east-1, us-west-2 as of 12/2016
  - Additional regions available*
- …but is not, by default, hosting across regions
  - Multiple regions has cost implications for ESDIS
  - NGAP favors explicit region architecture

- **Takeaway**: *Understand your users, plan your regions, and communicate them to NGAP*

EOSDIS

** There may be lag time for setup*

Every instance is three* instances

- NGAP uses a promotion model for all apps
  - SIT – developer testing beyond local machines
  - UAT – user acceptance testing, early access
  - Ops
- All applications must be functionally identical in each environment
  - An EC2 instance for a search engine in Ops means that same instance in lower environments
  - An application using 8 instances will require (at least) 24 instances in NGAP

- **Takeaway**: *Instance count matters!*

  (Also, see the next section…)

EOSDIS

** At a minimum!*

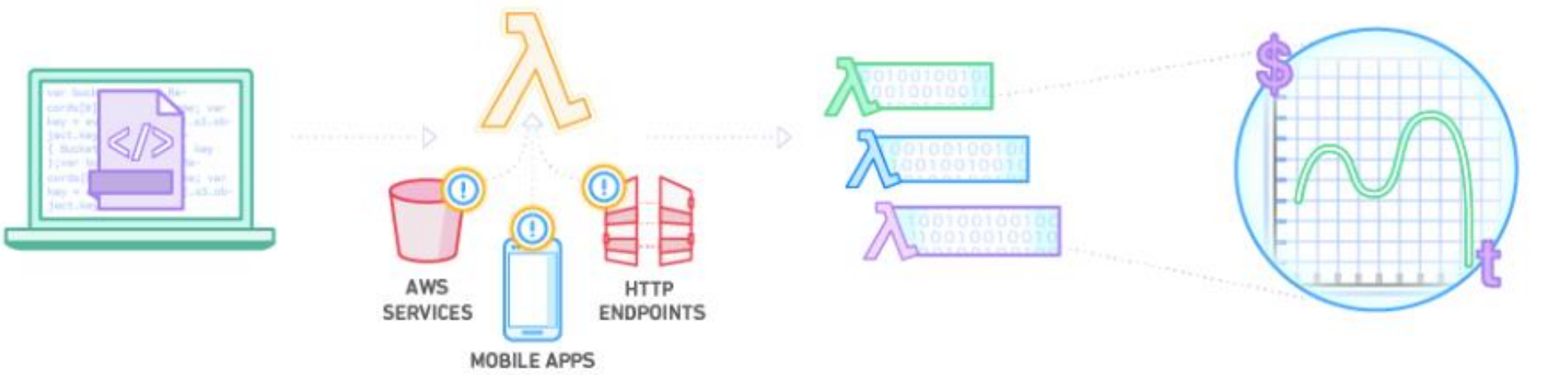
This is before considering…

- User behavior
- Staff cost savings
- Development cost savings
- Inter-region costs
- Data lifecycle modeling
- Application migration costs – both in and out
- Managing “consumption” based cost model

Cost Lesson 2

EXPLORE ALTERNATIVE ARCHITECTURES FOR POSSIBLE COST SAVINGS

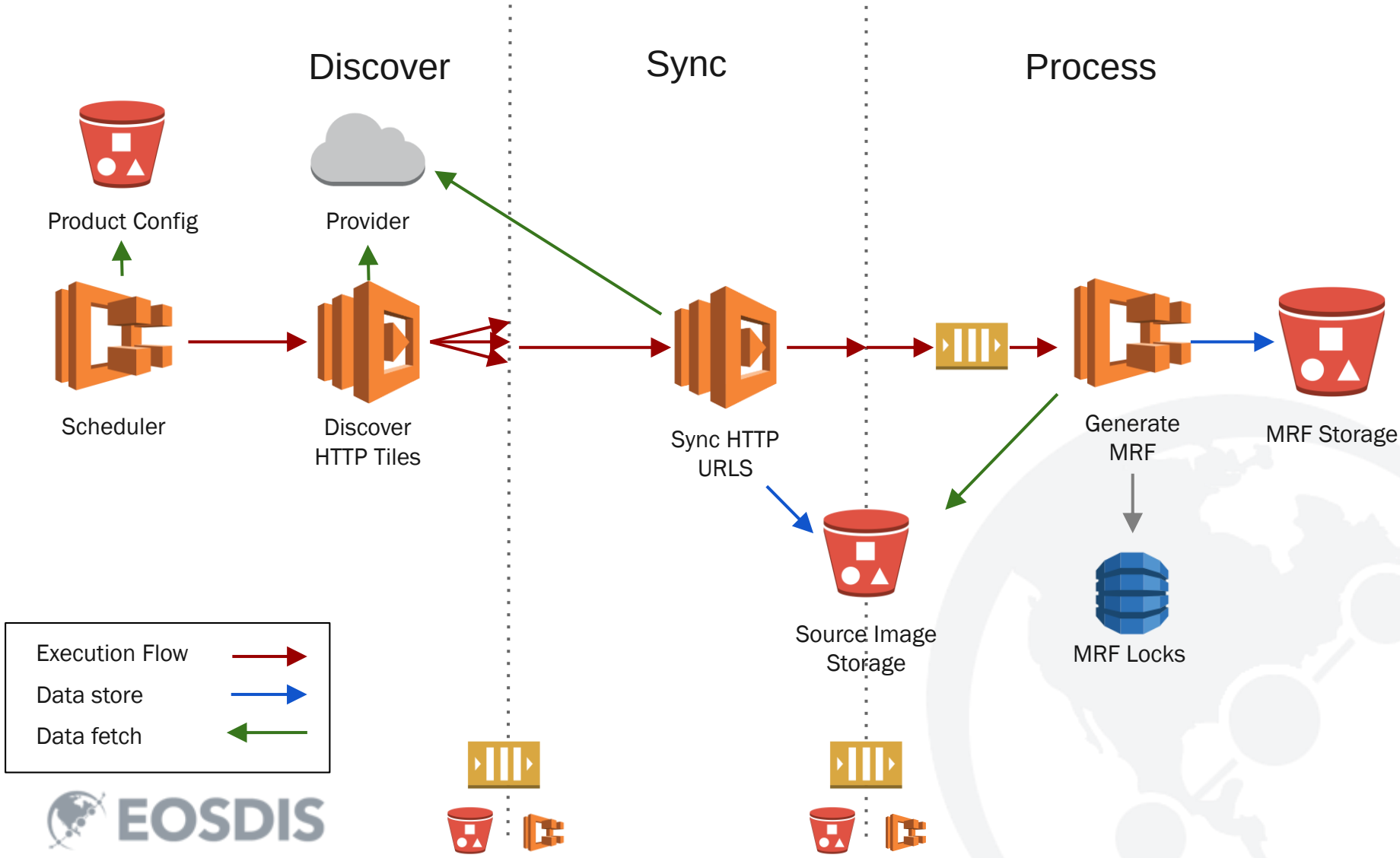
Use (just) what you need



EOSDIS

Graphic courtesy of http://amzn.to/11

Ingest: MODAPS Tiles



The Big 4… but serverless

1. ~~EC2 Instances~~
   - Zero to heavily reduced instances
2. ~~EBS Storage~~
   - Less EC2 generally means less EBS
3. Data Transfer
   - Notably: egress, egress, and egress
4. ELBs
   - More ELBs and more traffic = more cost

EOSDIS

** EBS storage has costs associated even when not in use by a running EC2 instance*

May not use for time purposes…

Need graphic for multiple egress models, requestor pays architecture, etc.

EOSDIS

Psycho-Social

GO HANDS-ON QUICKLY

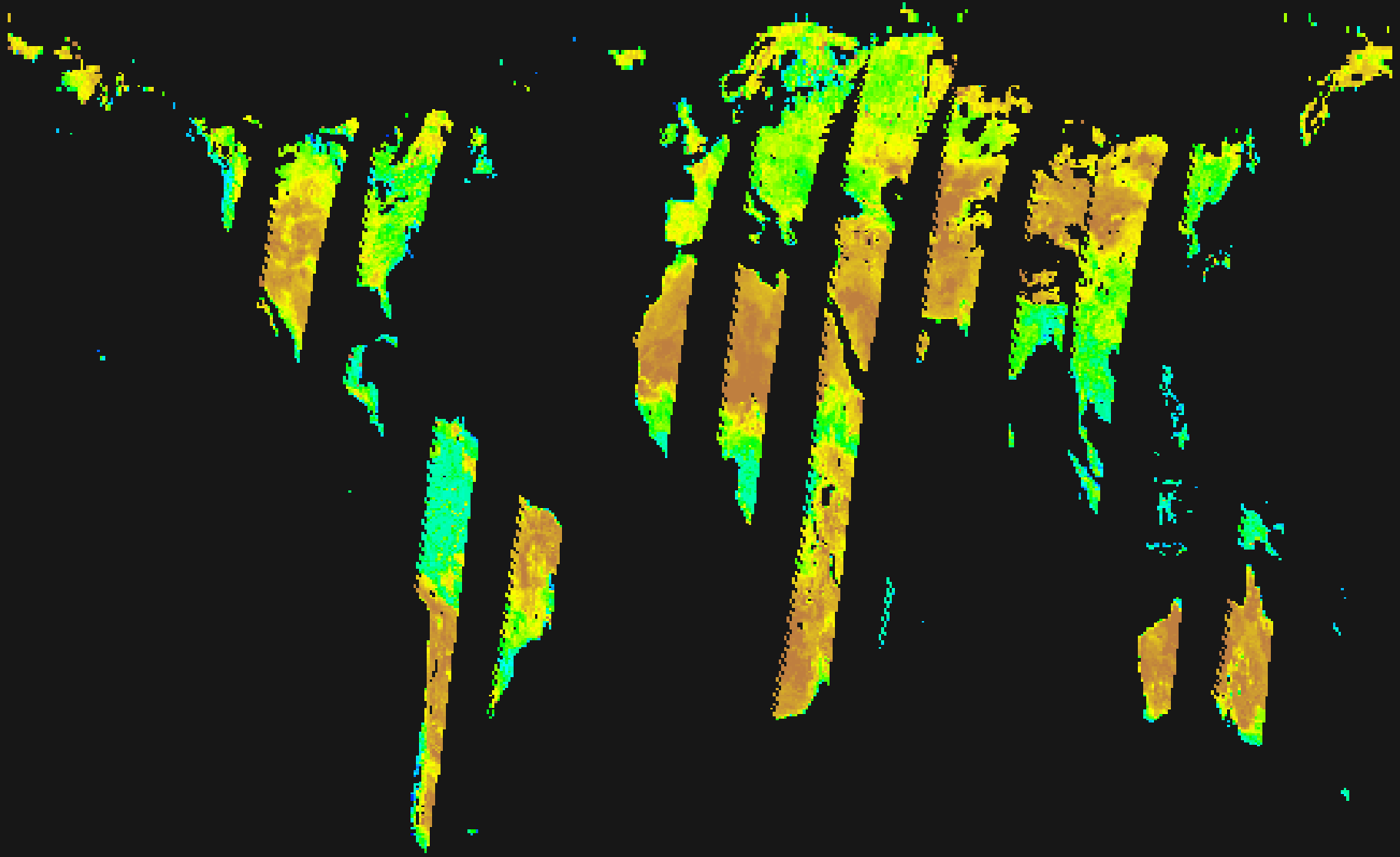

EED2
HACKFEST
2016
CAPABLE OF
SPENDING ENTIRE
ANNUAL BUDGET


EED2
HACKFEST
2016
MOST LIKELY TO
GO INTO
OPERATIONS


EED2
HACKFEST
2016
MOST
INNOVATIVE


EED2
HACKFEST
2016
MOST
POINTS

EOSDIS



Psycho-Social

UNDERSTAND THE OPERATIONS TEAM'S NEEDS

Current procedures may not translate directly

- Tailing / Grepping logs
- SSHing into machines to start / stop /restart services
- Monitoring specific hostnames
- Existing “operations” scripts
- Current dashboards vs AWS Console

Understand WHY they do what they do – you may need to find another way to do it.

Summary

- Enable cloud native architectures by strongly preferring cloud services
- AWS has very low internal latency, but trust nothing.
- Involve security from the very beginning
- Modeling TCO is extremely complicated
- Explore alternative architectures for possible cost savings
- Go hands-on quickly
- Incorporate Operations' Needs

Helpful Resources

- AWS Pricing: http://amzn.to/218Jr1G
- AWS Cost Optimization: http://amzn.to/2g3813l
- Decoding Your AWS Costs: http://bit.ly/1XCIzSk
- Minimizing AWS Transfer Costs: http://bit.ly/1njNOtJ
- Common Expensive Mistakes: http://bit.ly/1JR1NQb
- Serverless Architectures: http://amzn.to/1120t91

Questions?

Dan Pilone

dan@element84.com

This material is based upon work supported by the National Aeronautics and Space Administration under Contract Number **NNG15HZ39C.**

Raytheon

EOSDIS